The Bayes cost in the binary decision problem

Bernhard K. Meister*

Department of Physics, Renmin University of China, Beijing, China 100872

(Dated: October 25, 2011)

The problem of quantum state discrimination between two wave functions with fixed prior is considered. The optimal minimum error probability for the state discrimination is known to be given by the Helstrom bound. A new strategy is introduced here whereby a series of 'negative-result experiments' à la Renninger is carried out to modify the wave function. The resulting wave function is then probed using a conventional optimal strategy, and the error probability is calculated. It is shown that in some cases the Helstrom bound can be violated, i.e. the state discrimination can be realised with a smaller error probability.

I. INTRODUCTION

Experimental design and data analysis are common challenges in science, and particular acute in quantum mechanics, due to recurring questions about foundational issues and restrictions on measurements. In the literature many different approaches are discussed. For example, there is the information theoretic approach, where one maximizes the mutual information, and the minimax approach, where one minimizes the maximum cost of a set of strategies assuming a perfidious opponent. Both methods are popular, but this paper will instead use the Bayes procedure to minimize the expected cost, since the existence of a *prior* associated with the states to be distinguished is assumed.

Two disparate concepts are combined in the paper: quantum state discrimination and 'negative measurements'. Quantum state discrimination or more generally called Bayesian hypothesis testing was developed by Helstrom and others [4, 5, 9]. The particular problem of quantum state discrimination between two possible states with given prior and transition probability was studied in various settings, e.g. for a relatively recent paper on this topic see Brody *et al.* [2]. It is generally accepted, but will be challenged in the paper, that the optimal Bayes 0-1 cost in the binary case, is given by the Helstrom bound, which only depends on the prior and the transition probability between the states and can be written in a simple closed form. To avoid confusion a cost of zero is assumed in the case of a correct choice, and cost one, if the choice is incorrect.

Central to the paper and explained next is the concept of 'negative-result experiments', also called 'negative measurements'. These are measurements, where the possible absence of an observable event, leads to a modification of the wave function. The idea was introduced more than half a century ago by Renninger [8]. Simply put, if a quantum particle has a non-zero a priori probability in some regions in space, then a 'negative measurements' of one of these regions, e.g. a non-appearance of particle on a screen, has to be included in the calculation of the posterior distribution of the wave function. One prominent application is the quantum Zeno effect [7], i.e. repeated measurements with high survival probability due to shortness of the time interval between measurements, with an elegant realization in Kwiat *et al.* [6]. Some further interesting examples of 'negative measurements' can be found in Dicke [3], Bender *et al.* [1], and several related papers by the author easily found in the arixv.

Next a description of the setup and the procedure to be analyzed. With fixed probabilities, given by the prior, one of two quantum states is put into a quantum system governed by a specially chosen Hamiltonian H. Two strategies for calculating the Bayes cost are proposed. In the first strategy, the combination of prior and transition probability between the two quantum states alone is sufficient to calculate the conventional optimal minimum error probability, i.e. the Helstrom bound. This error probability can be achieved by a well-known measurement strategy [4] not further discussed in the paper. The second strategy for calculating the error probability is novel. It starts with an evolution of the system followed by a 'negative measurement'. The same procedure of evolution and measurement is carried out repeatedly. Finally the system is probed using a conventional measurement strategy to optimize the Bayes bound. For some specific quantum states the Helstrom bound can be higher than the new Bayes cost.

There are two main motivations for the work presented here. On the one hand it will possibly shed some light on foundational issues in quantum measurement theory, and on the other hand there are not many practical problems in quantum information theory, which do not in one way or another depend on optimal state discrimination.

The structure of the paper is as follows. In the second section the the binary choice problem for a particularly simple pair of examples states is tackled, and the strategy is described for breaching the Helstrom bound. In the conclusion the result is briefly restated and some general comments added.

*Electronic address: b.meister@imperial.ac.uk

II. BREAKING THE HELSTROM BOUND WITH A SEQUENCE OF 'NEGATIVE MEASUREMENTS' FOR A PAIR OF ALMOST ORTHOGONAL STATES

In this section it is shown how a series of 'negative measurements' in an appropriate direction can modify the state in a beneficial way and lead to a pair of states with larger distance and sufficiently high survival probability to challenge the Helstrom bound.

As part of the set up a 5-dimensional Hamiltonian is chosen of the form

$$H = \begin{pmatrix} E_0 & \delta & 0 & 0 & 0 \\ \delta & E_0 & 0 & 0 & 0 \\ 0 & 0 & E_0 & \delta & 0 \\ 0 & 0 & \delta & E_0 & 0 \\ 0 & 0 & 0 & 0 & E_1 \end{pmatrix}$$

with three distinct Eigenvalues, i.e. $E_0+\delta, E_0-\delta$ & E_1. The first and second Eigenvalue are each associated with two Eigenstates, e.g.

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

The last Eigenstate associated with the Eigenvalue E_1 is of the form $(0,0,0,0,1)$.

After describing some characteristics of Hamiltonian we now consider as an example a pair of particular simple test states to be distinguished. The two normalized candidate states have the form

$$\begin{pmatrix} a \\ 0 \\ 0 \\ 0 \\ b\,\delta \end{pmatrix} \& \begin{pmatrix} 0 \\ 0 \\ 0 \\ a \\ b\,\delta \end{pmatrix}.$$

with $a^2+(b\delta)^2=1$. The initial transition probability between the two states is $\cos^2(\alpha)=b^2\delta^2$, and the corresponding conventional minimal 0-1 Helstrom cost is

$$\frac{1}{2}\Big(1-\sqrt{1-4\xi(1-\xi)\cos^2(\alpha)}\Big).$$

In our particular case, we choose δ to be small and expand the cost and the wave functions in terms of this parameter. The prior ξ is fixed to be $1/2$, and as a consequence the cost can be written as

$$\frac{1}{2}\Big(1-\sqrt{1-4\xi(1-\xi)\cos^2(\alpha)}\Big)=\frac{1}{4}b^2\delta^2+O(\delta^4).$$

At the start of the process the two possible initial states evolve for a time Δt under the influence of the Hamiltonian H. To simplify matters we keep terms up to δ^2 and as a result the evolved state can be written either as

$$\begin{pmatrix} a\,e^{\imath E_0\Delta t}-\frac{1}{2}a\,(\Delta t)^2\delta^2 e^{\imath E_0\Delta t}+O(\delta^3) \\ \imath\,\delta\,a\,\Delta t\,e^{\imath E_0\Delta t}+O(\delta^3) \\ 0 \\ 0 \\ \delta\,b\,e^{\imath E_2\Delta t} \end{pmatrix}$$

or

$$\begin{pmatrix} 0 \\ 0 \\ \imath\,\delta\,a\,\Delta t\,e^{\imath E_0\Delta t}+O(\delta^3) \\ a\,e^{\imath E_0\Delta t}-\frac{1}{2}a\,(\Delta t)^2\delta^2 e^{\imath E_0\Delta t}+O(\delta^3) \\ \delta\,b\,e^{\imath E_2\Delta t} \end{pmatrix}.$$

Next a measurement is carried out in the $(0,1,1,0,0)$ direction. The survival probability for both candidate states is $1-\frac{1}{2}a^2\,\delta^2(\Delta t)^2+O(\delta^3)$. If the state 'survives', a normalization is necessary and the state becomes either

$$(1-\frac{1}{2}\,a^2\,\delta^2\,(\Delta t)^2)^{-1/2}\begin{pmatrix} a\,e^{\imath E_0\Delta t}-\frac{1}{2}a\,(\Delta t)^2\delta^2 e^{\imath E_0\Delta t} \\ \frac{1}{2}\imath\,\delta\,a\,t\,e^{\imath E_0\Delta t} \\ -\frac{1}{2}\imath\,\delta\,a\,t\,e^{\imath E_0\Delta t} \\ 0 \\ \delta\,b\,e^{\imath E_2\Delta t} \end{pmatrix}$$

or

$$(1-\frac{1}{2}\,a^2\,\delta^2\,(\Delta t)^2)^{-1/2}\begin{pmatrix} 0 \\ -\frac{1}{2}\imath\,\delta\,a\,\Delta t\,e^{\imath E_0\Delta t} \\ \frac{1}{2}\imath\,\delta\,a\,\Delta t\,e^{\imath E_0\Delta t} \\ a\,e^{\imath E_0\Delta t}-\frac{1}{2}a\,(\Delta t)^2\delta^2 e^{\imath E_0\Delta t} \\ \delta\,b\,e^{\imath E_2\Delta t} \end{pmatrix},$$

since a 'negative outcome' of the measurement is connected with the projection of the initial pre-measurement state onto a state orthogonal to the measurement direction such that the sum of the two orthogonal states recreates the pre-measurement state. Higher order terms involving δ^3 and beyond are ignored.

After carrying out the same procedure including the measurement k-times, one has with the survival probability either the state

$$(1-\frac{1}{2}a^2\,\delta^2\,(\Delta t)^2)^{-k/2}\begin{pmatrix} a\Big(1-\frac{1}{4}\,k(k+1)\,(\Delta t)^2\delta^2\Big)e^{\imath E_0\,k\,\Delta t} \\ \frac{1}{2}\imath\,k\,\delta\,a\,\Delta t\,e^{\imath E_0\,k\,\Delta t} \\ -\frac{1}{2}\imath\,k\,\delta a\,\Delta t\,e^{\imath E_0\,k\,\Delta t} \\ -\frac{1}{4}\,k(k-1)\,\delta^2 a\,(\Delta t)^2\,e^{\imath E_0\,k\,\Delta t} \\ \delta\,b\,e^{\imath E_2\,k\,\Delta t} \end{pmatrix}$$

or

$$(1-\frac{1}{2}a^2\,\delta^2\,(\Delta t)^2)^{-k/2}\begin{pmatrix} -\frac{1}{4}\,k(k-1)\,\delta^2 a\,(\Delta t)^2\,e^{\imath E_0\,k\,\Delta t} \\ -\frac{1}{2}\imath\,k\,\delta\,a\,\Delta t\,e^{\imath E_0\,k\,\Delta t} \\ \frac{1}{2}\imath\,k\,\delta\,a\,\Delta t\,e^{\imath E_0\,k\,\Delta t} \\ a\Big(1-\frac{1}{4}\,k(k+1)\,(\Delta t)^2\delta^2\Big)e^{\imath E_0\,k\,\Delta t} \\ \delta\,b\,e^{\imath E_2\,k\,\Delta t} \end{pmatrix}.$$

The survival probability slowly declines in each iteration and after k-iterations reduces to $1-\frac{1}{4}a^2\,k\,\delta^2\,(\Delta t)^2+$

$O(\delta^3)$. If the system 'survives' unscathed all the k-iterations the transition probability between the two candidate states evolves to $\Big(1+\frac{1}{4}a^2\,k\,\delta^2\,(\Delta t)^2\Big)\Big(b^2\delta^2-2\,a^2\,k^2\,\Delta t\,\delta^2\Big)+O(\delta^3)$. The product $2k\,a\,\Delta t$ will be equal to b, if the free parameters, namely the length of the time interval between measurements Δt and the number of iteration k are chosen appropriately. As a result the transition probability can be reduced to zero and one ends up with two orthogonal states after the k^{th}-measurement. It might seem strange that unlike in the standard Zeno-effect we choose δ as the small parameter instead of Δt, but this is just to simplify the example calculation and does not have any other significance. To avoid confusion, the whole process also works, if Δt is regarded as the small parameter, and the pair of initial states are $(a,0,0,0,\Delta t\,b)$ and $(0,0,0,a,-\Delta t\,b)$.

The Helstrom cost as a result of the new procedure can now be written as the sum of two parts, corresponding to the case where the state 'survives' all the k-iterations, and the case, where the state does not 'survive' the repeated measurements. In the first case, the Helstrom cost, due to the orthogonality of the resulting states, goes down to zero for our specific choice of parameters. In the second case the cost is $1/2$, i.e. the cost if zero information is available, times the probability of this alternative. The total new cost is therefore $k\,a^2\,(\Delta t)^2\,\delta^2/4+O(\delta^3)$, since in each of the k-iterations the probability of non-survival is $a^2\,(\Delta t)^2\,\delta^2/2+O(\delta^3)$.

After the preparatory calculations we can directly compare the old and the new cost, if we further set b equal to $2k\Delta t a$, as described above. As a consequence, one has either the original cost of the form $a^2\,k^2\,(\Delta t)^2\,\delta^2+O(\delta^4)$ or the new cost of $k\,a^2\,(\Delta t)^2\,\delta^2/4+O(\delta^3)$. One can directly see that this leads to a reduction in the cost, if k is larger or equal to 1, and the parameters a and Δt are of order 1. This counter-example destroys the general validity of the Helstrom bound.

III. CONCLUSION

The Helstrom bound in the binary quantum discrimination case can in principal be breached. This is the result of the paper. A counter-example to the general validity of the Helstrom bound has been provided.

Maybe it is worth pondering, why it is possible to break the Helstrom bound? The measurement process is non-linear and leads to this surprising behaviour of the cost function. In addition, the special relationship between amplitudes and probabilities exemplified in the role of k, i.e. the amplitude of the overlapping states increases linearly with k (and in conjunction the reduction in the survival probability is linear in k), but the transition probability between the candidate states increases quadratically in the same parameter.

As an aside, some limited similarities exist to an earlier paper by Bender *et al.* [1], where the insertion of a barrier and various types of 'negative measurements' led to an interesting entropy effect, and to other papers by the author.

In a companion paper to be presented at an upcoming conference a new optimal bound will be discussed and the number of steps required, relevant for quantum algorithms, to reach the new limit assessed. This bound will depend on the initial transition probability, the total evolution time, the temporal spacing between the measurements, as well as the relative energies of the basis states, and less directly on the implementation accuracy for the Hamiltonian.

As one can easily show the algorithm described not only works for almost orthogonal states, but in all cases, except some pathological exemptions, through the appropriate choice of the various degrees of freedom, i.e. the time between measurements Δt, the size of the off-diagonal term of the Hamiltonian δ, and the choice of the number of iterations k. Nevertheless even with this in-built flexibility there are limits to the improvements achievable beyond the Helstrom bound, which becomes more pronounced as the overlap between the initial states increase, e.g. two identical states can of course not be distinguished using the procedure described in this paper.

The author wishes to express his gratitude to D.C. Brody for stimulating discussions.

[1] Bender, C.M., Brody, D.C. & Meister, B.K., *Proc. of the Royal Society London* **A461**, 733-753 (2005).

[2] Brody, D.C. & Meister, B.K., *Phys. Rev. Lett.* **76** 1-5 (1996).

[3] Dicke, R.H. *Am. J. Phys.* **49**, 925 (1981).

[4] Helstrom, C.W., *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[5] Holevo, A.S., *Jour. Multivar. Anal.* **3**, 337 (1973).

[6] Kwiat, P.G., White, A.G., Mitchell, J.R., Nairz, O., Weihs, G., Weinfurter, H.,& Zeilinger, A. *Phys. Rev. Lett.* **83**, 4725 (1999).

[7] Misra, B. & Sundarshan, E.,C.,G. *J. Math. Phys.* **18**, 756 (1977).

[8] Renninger M., *Z. Phys.*, **136**, 251 (1953), and Renninger M., *Z. Phys.* **158**, 417 (1960).

[9] Yuen, H.P., Kennedy, R.S., and Lax, M., *IEEE Trans. Inform. Theory* **IT**-21, 125 (1975).